



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/704,039	11/01/2000	Christopher J. Brockett	M61.12-0213	2421
7590	11/17/2003		EXAMINER	
Westman Champlin & Kelly P A Suite 1600 International Centre 900 Second Avenue South Minneapolis, MN 55402-3319			PALMER, JUSTIN W	
			ART UNIT	PAPER NUMBER
			2654	<i>J</i>
DATE MAILED: 11/17/2003				

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary	Application No.	Applicant(s)
	09/704,039	BROCKETT ET AL.
	Examiner	Art Unit
	Justin Palmer	2654

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133).
- Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

- 1) Responsive to communication(s) filed on _____.
- 2a) This action is FINAL. 2b) This action is non-final.
- 3) Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

- 4) Claim(s) 1-24 is/are pending in the application.
- 4a) Of the above claim(s) _____ is/are withdrawn from consideration.
- 5) Claim(s) _____ is/are allowed.
- 6) Claim(s) 1-24 is/are rejected.
- 7) Claim(s) _____ is/are objected to.
- 8) Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

- 9) The specification is objected to by the Examiner.
- 10) The drawing(s) filed on 01 November 2000 is/are: a) accepted or b) objected to by the Examiner.
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).
- 11) The proposed drawing correction filed on _____ is: a) approved b) disapproved by the Examiner.
If approved, corrected drawings are required in reply to this Office action.
- 12) The oath or declaration is objected to by the Examiner.

Priority under 35 U.S.C. §§ 119 and 120

- 13) Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
 - a) All
 - b) Some *
 - c) None of:
 1. Certified copies of the priority documents have been received.
 2. Certified copies of the priority documents have been received in Application No. _____.
 3. Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).
- * See the attached detailed Office action for a list of the certified copies not received.
- 14) Acknowledgment is made of a claim for domestic priority under 35 U.S.C. § 119(e) (to a provisional application).
 - a) The translation of the foreign language provisional application has been received.
- 15) Acknowledgment is made of a claim for domestic priority under 35 U.S.C. §§ 120 and/or 121.

Attachment(s)

1) <input checked="" type="checkbox"/> Notice of References Cited (PTO-892)	4) <input type="checkbox"/> Interview Summary (PTO-413) Paper No(s). _____ .
2) <input type="checkbox"/> Notice of Draftsperson's Patent Drawing Review (PTO-948)	5) <input type="checkbox"/> Notice of Informal Patent Application (PTO-152)
3) <input checked="" type="checkbox"/> Information Disclosure Statement(s) (PTO-1449) Paper No(s) _____ .	6) <input type="checkbox"/> Other: _____ .

DETAILED ACTION

Claim Rejections - 35 USC § 102

1. The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

A person shall be entitled to a patent unless –

(b) the invention was patented or described in a printed publication in this or a foreign country or in public use or on sale in this country, more than one year prior to the date of application for patent in the United States.

2. Claims 1-8 are rejected under 35 U.S.C. 102(b) as being anticipated by Sproat et al., "A Stochastic Finite-State Word-Segmentation Algorithm for Chinese". *Computational Linguistics*, vol. 22:3, September 1996.

As per claim 1, Sproat et al. disclose a word segmentation method comprising:

identifying possible segments in the sequence of characters, at least two of the possible segments overlapping each other, using a "maximum matching method... one instance of which [Sproat et al. calls] the 'greedy algorithm'" (see method described in pp. 382-384, 393-394);

identifying an alternative sequence of characters for at least one of the possible segments, the alternative sequence of characters forming an alternative segment by choosing other possible characters that may also form a segment (see "maximum matching", greedy algorithm description cited above);

performing multiple syntactic analyses using the possible segments and the alternative segment, the analyses resulting in a full syntactic parse that utilizes and thereby results in a segmentation of the input sequence of characters (again please refer to the discussion of "maximum matching", greedy algorithm).

As per claims 2 and 3, Sproat et al. disclose a method according to claim 1 wherein:

performing multiple syntactic analyses comprises performing analyses that result in a parse containing the alternate segment (see description of "maximum matching", greedy algorithm);

identifying an alternative sequence of characters for a possible segment comprises identifying an alternative sequence of characters that has a different number of characters than the possible segment (“maximum matching”, as defined in the cited entry above will produce segments of varying character size).

As per **claim 5**, Sproat et al. disclose a method according to claim 1 wherein:

identifying an alternative sequence of characters comprises performing inflectional morphology on a possible segment by including “morphological rules” in order to expand the dictionary used in nonstochastic lexical-knowledge-based word segmentation methods (see Sproat et al., pp. 382-389).

As per **claim 6**, Sproat et al. disclose a method according to claim 1 wherein:

identifying an alternative sequence of characters comprises identifying orthographic variations of a possible segment (Sproat et al., p. 384), a fact elucidated by the discussion of orthographic words in languages such as Chinese, a language in which orthographic words are written with no spaces between them, and orthographic forms are identified through use of a dictionary (Sproat et al., pp. 378-379).

As per **claim 7**, Sproat et al. disclose a method according to claim 6 wherein:

identifying orthographic variations comprises identifying a preferred orthographic form for the possible segment (*ibid*, via dictionary lookup).

As per **claim 8**, Sproat et al. disclose a method according to claim 1 wherein:

identifying orthographic variations comprises identifying orthographic variants that use a script other than the script of the characters in the possible segment by looking in Roman and other alphabets (see Sproat et al., p. 384, bottom paragraph).

3. Claims 9-10, 12-14 rejected under 35 U.S.C. 102(b) as being anticipated by International Patent WO 98/08169 to Carus et al.

As per **claim 9**, Carus et al. teach a system for identifying syntax in a string of characters from a non-segmented language, the system comprising:

a word breaker that generates a collection of words from the string of characters, the collection of words comprising at least two words that are derived in part from the same character in the string of characters (see fig. 3, items 60, 62, 64, 76; p. 6, ll. 19-33; p. 29, l. 17-p. 30, l. 9; examples: pp. 32-35), the word breaker utilizing:

· a lexical record set (via dictionary lookups) that is used to derive words from the collection of words by taking the words directly from the string of characters (see p. 3, l. 23; p. 5, ll. 6-18; p. 6, ll. 19-33; p. 36 l. 28-p. 37 l. 10) through a database analysis module and a heuristic analysis module and;

· a variants constructor that is used to derive word variants of words found in the string of characters, each word variant being added to the collection of words and having a different sequence of characters than the sequence of characters associated with the word in the string of characters from which it is derived (see previous cite, descriptions of heuristic and database analysis modules; p. 36, l. 14-p. 36, l. 10; fig. 5, step 100; fig. 6); and

· a syntax parser that performs a syntactic analysis using the collection of words produced by the word breaker to produce a syntax parse; the process outlined in Carus et al. therein comprising a syntax parse which identifies syntactic units of a given corpus per applicant's claim and specification (see fig. 2; p. 9, ll. 22-36; p. 10, ll. 1-9).

As per **claim 10**, Carus et al. teach a system according to claim 9 wherein:

the variants constructor comprises inflectional morphology rules (see previous site under variants constructor, claim 9).

As per **claim 12**, Carus et al. teach a system according to claim 9 wherein:

Art Unit: 2654

the variants constructor comprises an orthographic variants structure that indicates the orthographic variants of words found in the string of characters (see previous cites in claim 9 regarding the lexical record set and variants constructor), a dictionary being used to check character orthography.

As per **claim 13**, Carus et al. teach a system according to claim 9 wherein:

at least one word variant has a different number of characters than the word from which it is derived, the variants constructor disclosed in Carus necessarily creating a variant with a different number of characters (see previous cite in claim 9 regarding variants constructor).

As per **claim 14**, Carus et al. teach a system according to claim 9 wherein:

at least one word variant includes a character that is not present in the string of characters, the variants constructor disclosed in Carus necessarily creating a variant with alternate characters (see previous cite in claim 9 regarding variants constructor).

Claim Rejections - 35 USC § 103

4. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

5. This application currently names joint inventors. In considering patentability of the claims under 35 U.S.C. 103(a), the examiner presumes that the subject matter of the various claims was commonly owned at the time any inventions covered therein were made absent any evidence to the contrary. Applicant is advised of the obligation under 37 CFR 1.56 to point out the inventor and invention dates of each claim that was not commonly owned at the time a later invention was made in order for the examiner to consider the applicability of 35 U.S.C. 103(c) and potential 35 U.S.C. 102(e), (f) or (g) prior art under 35 U.S.C. 103(a).

Art Unit: 2654

6. Claim 11 is rejected under 35 U.S.C. 103(a) as being unpatentable over Carus et al. in view of Sproat et al.

As per **claim 11**, Carus et al. disclose the teachings of claim 10, on which claim 11 depends (see cite in claim 10 rejection above).

Carus et al. fail to explicitly teach inflectional morphology rules that are capable of identifying a word's lemma from its inflectional form in the string of characters.

Sproat et al. disclose inflectional morphology rules that can identify a word's stem from inflectional morphology rules (see pp. 386, 398).

At the time of the invention, it would have been obvious for one skilled in the art to modify Carus et al. to include inflectional morphology rules capable of identifying a word's stem from an inflectional form because Sproat et al. teach the use of inflectional morphology rules in order to better segment an unsegmented character string (see pp. 386, 398).

7. Claims 15-24 are rejected under 35 U.S.C. 103(a) as being unpatentable over Carus et al. in view of Sproat et al.

As per **claims 15-24**, Carus et al. and Sproat et al. disclose methods according to claims 1-8.

Carus et al. and Sproat et al. do not specifically disclose a computer system and a computer-readable medium to execute said methods.

At the time of invention, it would have been obvious to one skilled in the art to modify Carus et al. and Sproat et al. to include a computer system to execute the disclosed methods and a computer-readable medium coupled to said system; such modification is implicit in the assumptions and spirit of both inventions.

Art Unit: 2654

Such a modification would have been obvious in order to execute said methods, update the method and executable code, provide read-write storage for the executable code and data, evaluate its performance, and incorporate the system and media into a product.

Conclusion

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Justin Palmer whose telephone number is (703) 305-8663. The examiner can normally be reached on Monday-Thursday 7:00 AM-5:15 PM Eastern.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Talivaldis Ivars Smits can be reached on (703) 306-3011. The fax phone number for the organization where this application or proceeding is assigned is (703) 872-9306.

Any inquiry of a general nature or relating to the status of this application or proceeding should be directed to the receptionist whose telephone number is (703) 305-4700.

JP
22 October 2003



RICHMOND DORVIL
SUPERVISORY PATENT EXAMINER